

Supplementary Material

Alpha Helix Nucleation by a Simple Cyclic Tetrapeptide

Huy N. Hoang,^{A,B} Chongyang Wu,^{A,B} Renee L. Beyer,^A Timothy A. Hill,^{A,C} and David P. Fairlie^{A,C}

^ADivision of Chemistry and Structural Biology and ARC Centre of Excellence in Advanced Molecular Imaging, Institute for Molecular Bioscience, The University of Queensland, Brisbane, Qld 4072, Australia.

^BThese authors contributed equally to this paper.

^CCorresponding authors. Email: t.hill@imb.uq.edu.au; d.fairlie@imb.uq.edu.au

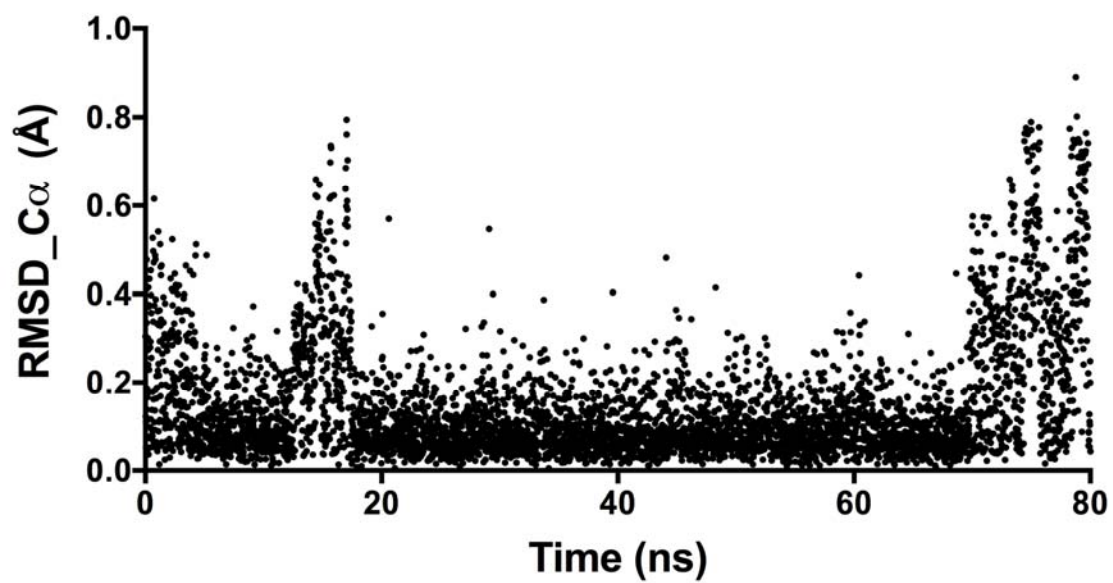


Figure S1. Molecular dynamics simulations for **4**, backbone RMSD variations over 80 ns.

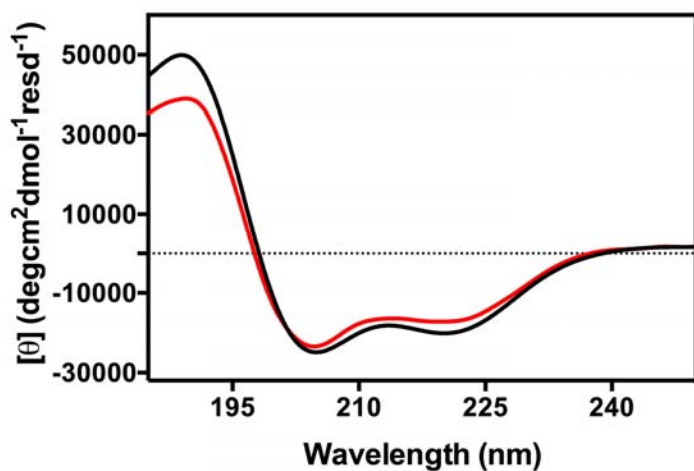


Figure S2. CD spectra for peptides **9** and **10** (50 μ M) in 10 mM phosphate buffer (pH 7.2, 298K); **9** (black, 69% helicity) and **10** (red, 60% helicity).

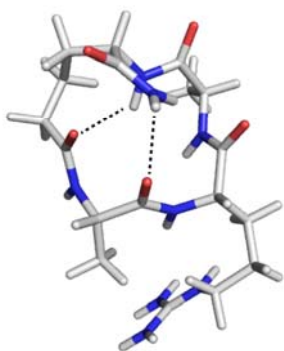


Figure S3. Solution NMR calculated structure for **4** indicating two hydrogen bonds (dotted lines) derived from amide proton temperature coefficients.

EXPERIMENTAL PROCEDURES

Abbreviations

DIPEA, diisopropylethylamine; DMF, dimethylformamide; ESI-MS, electrospray ionization mass spectrometry; Fmoc, 9-fluorenylmethyloxycarbonyl; HBTU, 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyl uronium hexafluorophosphate; HCTU, 2-(1*H*-6-chlorobenzotriazol-1-yl)-1,1,3,3-tetramethyl uronium hexafluorophosphate; HR-MS, High-resolution mass spectroscopy; MBHA, 4-methyl-benzylhydramine; Mmt, 4-methoxyltrityl; Mtt, 4-methyltrityl; OPip, phenyl isopropyl ester; PyBOP, benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate; RP-HPLC, reserved-phase high performance liquid chromatography; UHPLC, reserved-phase ultra high performance liquid chromatography; RT, room temperature; SPPS, solid-phase peptide synthesis; *t*Bu, *tert*-butyl; TFA, trifluoroacetic acid; TFE, 2,2,2-trifluoroethanol; TIS, triisopropylsilane; Trt, triphenylmethyl; Pd(PPh₃)₄, Tetrakis(triphenylphosphine)palladium.

Materials

All solvents and reagents used during peptide chain assembly were peptide synthesis grade and purchased from commercial suppliers unless otherwise stated. N^α-Fmoc-L- α -amino adipic acid- δ -allyl ester was prepared based the reported procedure.^[1]

[1] Long, Y.-Q.; Lung, F.-D. T.; Roller, P. P., Global optimization of conformational constraint on non-phosphorylated cyclic peptide antagonists of the Grb2-SH2 domain. *Bioorg. Med. Chem.* **2003**, *11* (18), 3929-3936

Synthesis of compound 1:

Compound **1** was synthesized and characterized as previously reported^[2]

[2] Shepherd, N. E.; Hoang, H. N.; Abbenante, G.; Fairlie, D. P., Single Turn Peptide Alpha Helices with Exceptional Stability in Water. *J. Am. Chem. Soc.* **2005**, *127* (9), 2974-2983.

Synthesis of compound 2

Compound **2** was synthesized and characterized as previously reported^[3]

[3] Hoang, H. N.; Driver, R. W.; Beyer, R. L.; Malde, A. K.; Le, G. T.; Abbenante, G.; Mark, A. E.; Fairlie, D. P., Protein α -Turns Recreated in Structurally Stable Small Molecules. *Angew. Chem., Int. Ed.* **2011**, *50* (47), 11107-11111.

Synthesis of compound 3:

Compound **3** was synthesized and characterized as previously reported^[3]

[3] Hoang, H. N.; Driver, R. W.; Beyer, R. L.; Malde, A. K.; Le, G. T.; Abbenante, G.; Mark, A. E.; Fairlie, D. P., Protein α -Turns Recreated in Structurally Stable Small Molecules. *Angew. Chem., Int. Ed.* **2011**, *50* (47), 11107-11111.

Synthesis and characterization of compound 4:

Compound **4** was synthesized manually by standard solid phase methods using HCTU/DIPEA activation for Fmoc chemistry on Rink Amide MBHA resin (substitution 0.70 mmol.g⁻¹, 0.70 mmol scale syntheses, 1.00 g resin). Four equivalents of Fmoc-protected amino acids, 4 equivalents of HCTU, and 4.8 equivalents of DIPEA were employed in each coupling. Fmoc deprotections were achieved by 2 × 3 min. treatments with excess 1:1 piperidine:DMF. Coupling yields were monitored by quantitative ninhydrin assay. The peptide was cleaved from the resin and protecting groups simultaneously removed by treatment for 4h at 22°C with a 10mL solution containing 95% trifluoroacetic acid (TFA): 2.5% H₂O: 2.5% triisopropylsilane (TIPS). TFA solutions were filtered, concentrated *in vacuo*, diluted with 50% A : 50% B then lyophilised. The linear peptide was purified by semi-preparative rp-HPLC on a 22 × 250 mm column using a linear gradient (0%B to 100%B)(30 min). Cyclisation of the side chain carboxylic acid to the free amino N-terminus was achieved by dissolving the linear peptide in DMF (8mM) and stirring with 1.2 equivalents of DPPA (diphenylphosphoroazide) and 2 equivalents of DIPEA at 4°C with the pH adjusted to pH 8 for 6 days in a dark environment. DMF was removed under reduced pressure prior to purification.

For the linear peptide **4a**, purification was achieved by semi-preparative rp-HPLC on the 22 x 250 mm column using a linear gradient (0%B to 100%B)(30 min) [R_t = 5.6

min]. Yield = 60%. Anal. rp-HPLC (linear gradient from 0% to 100% B over 30 min.): R_t = 12.31 min. MS: [obs. (M+H⁺) 459.20; calcd 459.27]. Purification of the cyclic peptide **4** was achieved by semi-preparative rp-HPLC on the 22 x 250 mm column using a complex gradient, isocratic (100%A)(5 min), isocratic (5%B)(5 min), then linear gradient (5%B to 100%B)(20 min), [R_t = 15.1 min]. Yield = 12%. Anal. rp-HPLC (linear gradient from 0% to 100% B over 30 min.): R_t = 13.39 min. MS: [obs. (M+H⁺) 441.2535; calcd 441.2568]. ¹H NMR (H₂O:D₂O) δ 8.60 (d, *J* 6.5 Hz, 1H, Ala1NH), 8.51 (d, *J* 6.5 Hz, 1H, Ala3NH), 7.88 (d, *J* 6.2 Hz, 1H, Arg2NH), 7.58 (d, *J* 7.4 Hz, 1H, homoglu4NH), 7.31 (s, 1H, NH₂), 7.23 (t, *J* 5.1 Hz, 1H, Arg2 ϵ NH), 7.11 (s, 1H, NH₂), 4.22 (m, 1H, Ala3H α), 4.21 (m, 1H, Arg2H α), 4.10 (m, 1H, Ala1H α), 4.10 (m, 1H, homoglu4H α), 3.22 (q, 2H, Arg2H δ), 2.38 (m, 1H, homoglu4H δ), 2.13 (m, 1H, homoglu4H δ), 1.99 (m, 1H, homoglu4H γ), 1.83 (m, 1H, homoglu4H γ), 1.82 (m, 2H, Arg2H β), 1.67 (m, 1H, Arg2H γ), 1.64 (m, 1H, homoglu4H β), 1.61 (m, 1H, Arg2H γ), 1.49 (d, *J* 7.2 Hz, 3H, Ala1H β), 1.45 (m, 1H, homoglu4H β), 1.42 (d, *J* 7.2 Hz, 3H, Ala3H β).

General automated SPPS for the synthesis of peptides

Peptides **5-15** were assembled on a peptide synthesizer (Symphony) using Rink Amide MBHA resin with 5 equiv. of Fmoc-protected amino acids, 5 equiv. of HCTU and 5 equiv. of DIPEA. Each coupling process involved 2 coupling cycles (30min x 2). Fmoc deprotection was achieved by 5 x 2 min treatments with excess 20% piperidine in DMF.

For peptides **8**, **11** and **13**, the *N*-terminus was acetylated with Ac₂O:DIPEA (0.87:0.47:15 mL) in DMF for 10 minutes.

General methods for the synthesis of cyclic peptides in solution phase 5-10

The linear peptides with free *N*-terminal amine were dissolved in DMF in a 50 mL round flask, the final concentrations about 0.001M, followed by adding 4.0 equivalent HATU and 4.0 equivalent DIPEA. Coupling progresses were monitored by ninhydrin test and ESI-MS. After the completion of the reaction, DMF was removed and then the crude residue was purified by RP-HPLC using a Phenomenex Luna C18 column eluting at a flow rate of 20 mL/min and a gradient of 0 to 40% buffer B (90% CH₃CN/10% H₂O/0.1% TFA) in 40 minutes and then lyophilised.

General methods for the synthesis of cyclic peptides on solid phase

For the cyclic peptides 12 and 14, the linear precursors were selectively deprotected by Pd(PPh₃)₄ (20% equiv) in DCM overnight on resin, then washed with DMF, DCM, sodium diethyldithiocarbamate/DMF (5 mg/10 mL; 20 min), DMF, and DCM until the dark colour disappear. The N-terminal cyclization of the resin bound peptide was performed with PyBOP (4.0 equiv) and DIPEA (4.0 equiv) in DMF at r.t overnight.

For cyclic peptides 15, Fmoc-Lys(Mtt)-OH and Fmoc-Asp(OPip)-OH were employed for incorporation of non-standard amino acids at positions 1 and 5 respectively.^[4] The peptide was assembled on the solid support as described above. The resin was then washed with DCM and treated repeatedly with 3% TFA in DCM (5 x 6 min). After washing with DMF, a solution of PyBOP (4.0 equiv) and DIPEA (4.0 equiv) in DMF was added to the resin and the reaction was agitated overnight.

[4] Hoang, H. N.; Driver, R. W.; Beyer, R. L.; Hill, T. A.; Aline, D. d. A.; Plisson, F.; Harrison, R. S.; Goedecke, L.; Shepherd, N. E.; Fairlie, D. P., Helix Nucleation by the Smallest Known α -Helix in Water. *Angew. Chem. Int. Ed. Engl.* **2016**, *55* (29), 8275-8279.

Cleavage from solid support and peptide purification

Peptides were cleaved from the resin by treatment with TFA:TIS:H₂O (95:2.5:2.5) for 2.5 h. The crude peptides were precipitated and washed with cold Et₂O, redissolved in solution containing 50% acetonitrile, 50% H₂O and 0.01% TFA and then lyophilised. Peptides were purified by RP-HPLC using a Phenomenex Luna C18 column eluting at a flow rate of 20 mL/min and a gradient of 0 to 40% buffer B (90% CH₃CN/10% H₂O/0.1% TFA) in buffer A (0.1% TFA in water) in 40 minutes.

Analytical methods

Analytical RP-HPLC was performed on an Agilent system, using a Phenomenex Luna C18 5 μ m (250 x 4.60 mm) column eluting at a flow rate of 1.0 mL/min and a gradient of 0 to 80% or 100 % buffer B (90% CH₃CN/10% H₂O/0.1% TFA in buffer A, 0.1% TFA in water) over 20 minutes.

Circular Dichroism Spectroscopy

Peptide solutions for CD were prepared from aqueous peptide stock solutions (540 μ L H₂O and 60 μ L D₂O) and the accurate molecular concentrations were determined by NMR. The final concentration of the peptide samples was 50 μ M in 10mM phosphate buffer pH 7.2. CD measurements were performed using a Jasco model J-710 spectropolarimeter which was routinely calibrated with

(1S)-(+)-10-camphorsulfonic acid. Spectra were recorded at room temperature (298K), with a 0.1 cm Jasco quartz cell over the wavelength range 260-185 nm at 50 nm/min, with a bandwidth of 1.0 nm, response time of 1 s, resolution step width of 1 nm and sensitivity of 20-50 Mdeg. Each spectrum represents the average of 3 scans. Spectra were analysed using the spectral analysis software and smoothed using 'adaptive smoothing' function. Concentrations were determined using the PULCON method.⁶ NMR solutions were prepared with 550 μ L of stock solution. 90° pulses were accurately determined and then 1D Spectra were acquired using the standard watergate sequence with a ns= 32-64, d1= 25-35s. Spectra were also acquired for a 4.76 mM solution of L-histidine as the reference standard. The fully resolved, most downfield amide resonance was integrated and used to calculate the concentration from the equation:

$$c_u = c_R \frac{S_U T_U \theta_{360}^U n_R r g_R}{S_R T_R \theta_{360}^R n_U r g_U}$$

where c is the concentration, S is the integral(in absolute units)/number of protons, T is the temperature in Kelvin, θ_{360} is the 360° rf pulse, n is the number of scans, and rg is the receiver gain used for measuring the reference (R) and unknown (U) samples.

Percentage Helicity of peptides were calculated from residue-molar ellipticity at 215 nm (for **1-4**) and at 222 nm (for longer peptides **5-15**) using the following equation:

Where $[\theta]_{\max}$ ($[\theta]_{\max} = [\theta]_{\infty}(n - x)/n$) is the maximum theoretical mean residue ellipticity for a helix of n residues, $[\theta]_{\infty}$ is the mean residue ellipticity of an infinite helix, and x is an empirical constant that can be interpreted as the effective number of amides missing as a result of end effects, usually about 2.4-4 (we used $x=3$) and $[\theta]_{\infty} = (-44000 + 250T)$ (T is temperature of the peptide solution in °C). $[\theta]_0$ is the mean residue ellipticity of the peptide in random coil conformation and equals to $(2220 - 53T)$ and $[\theta]_{222}$ ($[\theta]_{222} = 1/n \cdot [\theta_{\text{obs}}]/(10 \times l \times C)$) is the observed residue ellipticity of peptide at 222 nm. Where θ_{obs} = measured ellipticity in mdeg; n = number of peptide residues; C = sample concentration (mol/L); l = optical path length of the cell in cm.⁷

Proton NMR Spectroscopy

The samples for the NMR analyses of peptides were prepared by dissolving the peptide (1.5 to 2.2 mg) in 540 μ L H₂O and 60 μ L D₂O at pH 5.0. 1D and 2D ¹H-NMR

spectra were recorded on a Bruker CryoProbe Avance III 600 MHz spectrometer. 2D ^1H -spectra were recorded in phase-sensitive mode using time-proportional phase incrementation for quadrature detection in the t_1 dimension. The 2D experiments included TOCSY (standard Bruker mlevgpqh pulse program), ROESY (standard Bruker roesygpqh pulse program), NOESY (standard Bruker noesygpqh pulse program) and dqfCOSY (standard Bruker dqfcosygpqh pulse program). TOCSY spectra were acquired over 6887 Hz with 4096 complex data points in F_2 , 512 increments in F_1 and 32 scans per increment. ROESY and NOESY spectra were acquired over 6887 Hz with 4096 complex data points in F_2 , 512 increments in F_1 and 32 scans per increment. TOCSY, ROESY and NOESY spectra were acquired with several isotropic mixing times of 80, 100 ms for TOCSY; 200, 350 ms for ROESY and 150, 250 ms for NOESY. For all NMR experiments, water suppression was achieved using modified WATERGATE. For 1D ^1H NMR spectra acquired in $\text{H}_2\text{O}/\text{D}_2\text{O}$ (9:1), the water resonance was suppressed by low power irradiation during the relaxation delay (1.5 to 3.0 s). The variable temperature NMR experiments were performed over the range of 288-318K. Spectra were processed using Topspin (Bruker, Germany) software and ROE intensities were collected manually. The t_1 dimensions of all 2D spectra were zero-filled to 1024 real data points with 90° phase-shifted QSINE bell window functions applied in both dimensions followed by Fourier transformation and fifth order polynomial baseline correction. ^1H chemical shifts were referenced to DSS (δ 0.00 ppm) in water. $^3J_{\text{NHCH}\alpha}$ coupling constants were measured from 1D ^1H NMR and dqf-COSY spectra using XPLOR program.

Structure Calculations

The distance restraints used in calculating the structure for **4** and **7** in water were derived from ROESY and NOESY spectra (recorded at 298K) using mixing time of 350ms and 250 ms with 23 and 51 ROEs for **3** and **6** respectively. ROE cross-peak volumes were classified manually as strong (upper distance constraint $\leq 2.7\text{\AA}$), medium ($\leq 3.5\text{\AA}$), weak ($\leq 5.0\text{\AA}$) and very weak ($\leq 6.0\text{\AA}$). Standard pseudoatom distance corrections were applied for non-stereospecifically assigned protons ($\text{CH}_2 = 2\text{\AA}$ and $\text{CH}_3 = 1.5\text{\AA}$). To address the possibility of conformational averaging, intensities were classified conservatively and only upper distance limits were included in the calculations to allow the largest possible number of conformers to fit the experimental data. Backbone dihedral angle restraints were inferred from $^3J_{\text{NHCH}\alpha}$ coupling constants in 1D spectra, ϕ was restrained to $-65 \pm 30^\circ$ for $^3J_{\text{NHCH}\alpha} \leq 6\text{Hz}$ and to $-120 \pm 30^\circ$ for $^3J_{\text{NHCH}\alpha} \geq 8\text{Hz}$. There was clearly no evidence at all for *cis*-amides about peptide bonds (i.e. no $\text{CH}\alpha\text{-CH}\alpha$ ($i, i+1$) ROEs) in the ROESY spectra (in both 9:1 $\text{H}_2\text{O}/\text{D}_2\text{O}$ and 100% D_2O) so all ψ -angles were set to *trans* ($\psi = 180^\circ$).

Starting structures with randomised ϕ and ψ angles and extended side chains were generated using an *ab initio* simulated annealing protocol. The calculations were performed using the standard forcefield parameter set (PARALLHDG5.2.PRO) and topology file (TOPALLHDG5.2.PRO) in XPLOR-NIH with in house modifications to generated topology for homo-glutamic residue and cyclisation from side chain to back bone amide bond. Refinement of structures was achieved using the conjugate gradient Powell algorithm with 4000 cycles of energy minimisation and a refined forcefield based on the program CHARMM. Structures were visualised with Pymol and analysed for distance ($>0.2\text{\AA}$) and dihedral angle ($>5^\circ$) violations using noe.inp files. Final structures contained no distance violations ($>0.2\text{\AA}$) or angle violations ($>5^\circ$).

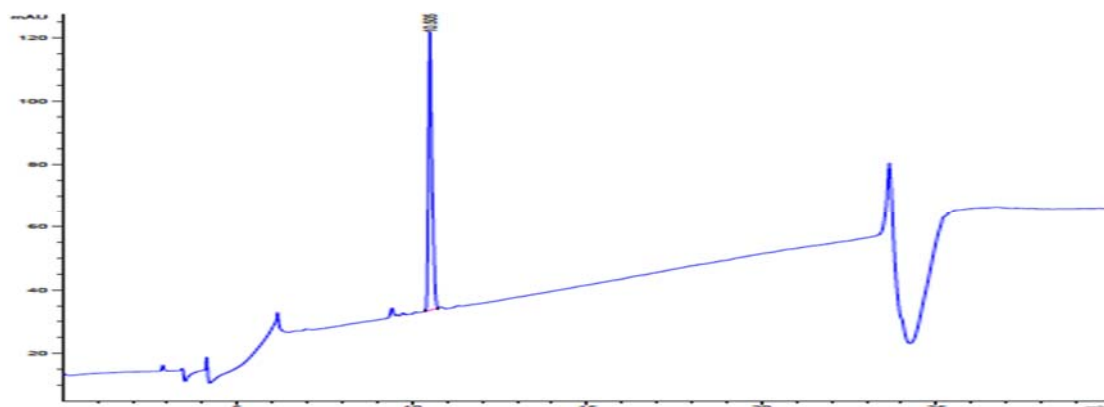
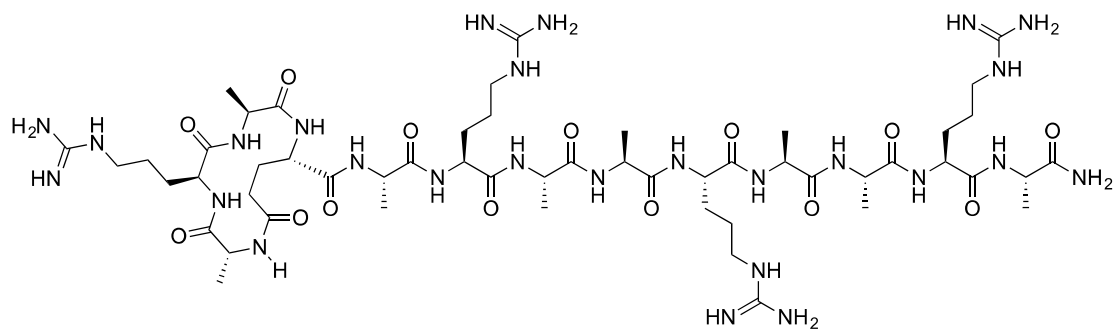
Molecular Dynamics (MD) Simulations

All simulations were performed using Maetro simulation package. The force field topologies of the cyclic peptide **4** were derived from the OP3 parameter set. The cyclic peptide was placed in a pre-equilibrated truncated octahedral box filled with 1038 SPC water molecules. The simulation was carried out at 298 K. After a steepest descent minimization step, the cyclic peptides were subjected to 100 ps of simulation with position restraints on the peptide to relax the solvent. The systems were subsequently equilibrated for 1 ns without restraints. The simulations were performed at constant temperature (298 K) and pressure (1 atm). This was achieved using a Berendsen thermostat with a coupling time of 0.1 ps and a Berendsen barostat with a coupling time of 0.5 ps. Nonbonded interactions were calculated using a twin-range cut-off. Interactions within the short-range cut-off of 0.8 nm were updated every time step. Interactions within the longer-range cut-off of 1.4 nm were updated every 5 time steps together with the pairlist. All bonds were constrained using the SHAKE algorithm with a geometric tolerance of 0.0001. Initial velocities were taken from the Maxwell-Boltzmann distribution at 298 K and was acquired for 80 ns. Backbone ($C\alpha$) atom-positional RMSD values were calculated after translational superposition of centres of mass and least-squares rotational fitting of atomic positions.

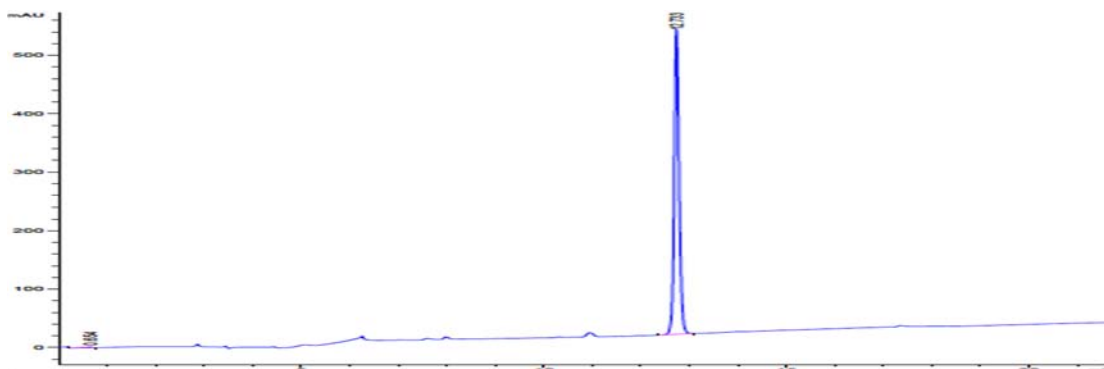
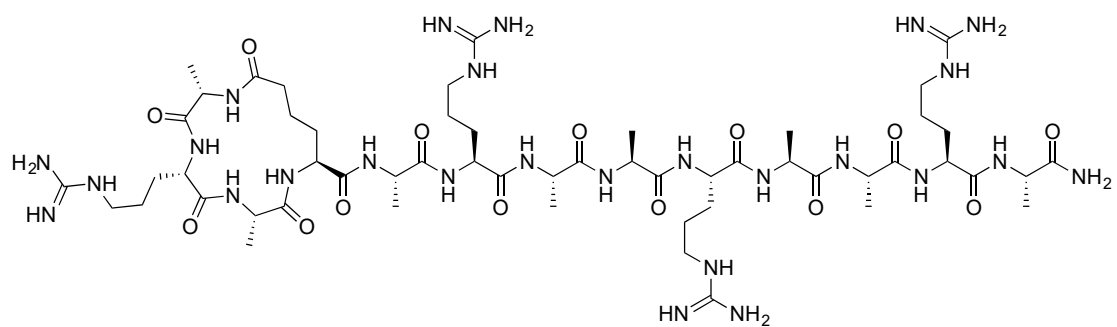
Analytical data (MS)

#	Exact Mass	Calculated Isotopic ions m/z [M+nH] ⁺ (n)	Isotopic ions m/z _{obs} [M+nH] ⁺ (n)
4	441.3	441.2568	441.2535
4a	459.27	459.27	459.20
5	1320.8	441.3, 331.9	441.2, 331.2
6	1320.8	661.8, 441.5, 331.2	661.3, 441.2, 331.2
7	1334.8	668.77, 446.18	669.46, 446.69
8	1394.8	466.20, 698.8	466.0, 698.7
9	1235.7	619.2, 413.1	619.1, 413.0
10	1249.7	626.2, 417.8	625.9, 417.6
11	955.5	956.1	955.7
12	1038.6	1039.2	1039.8
13	1476.8	14477.7	1478.8
14	1559.8	1560.7, 781.4	1561.4, 780.9
15	1701.9	852.5	851.9

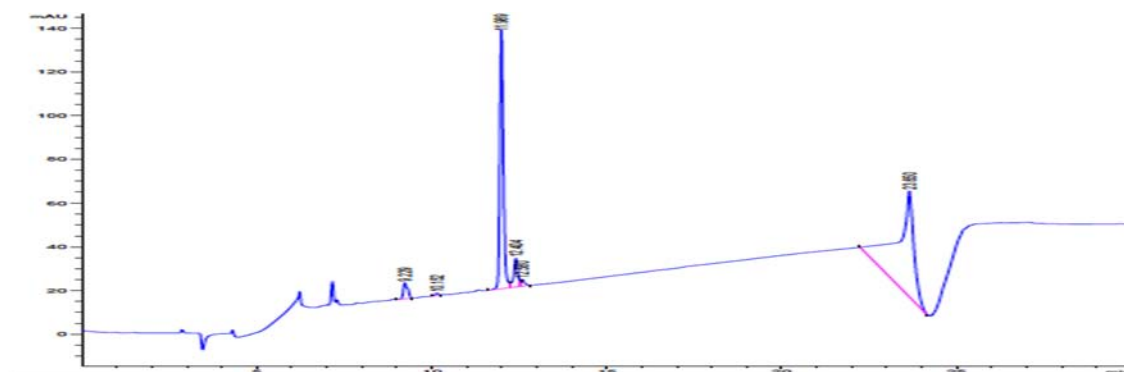
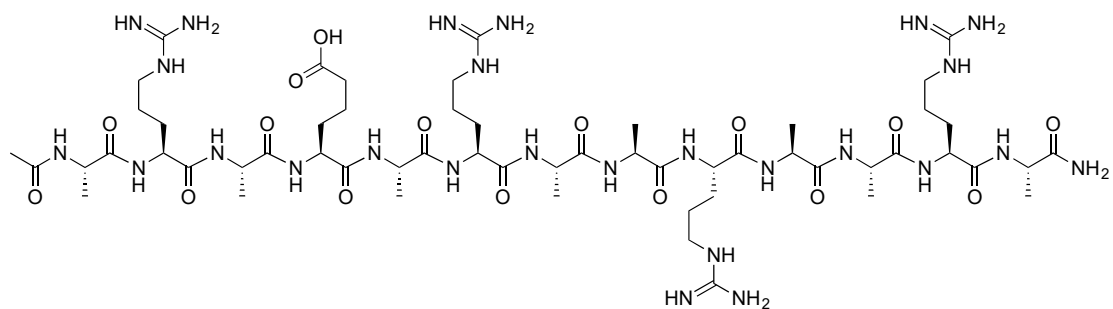
6

cyclo-(1,4)-[ARAE]ARAARA-NH₂

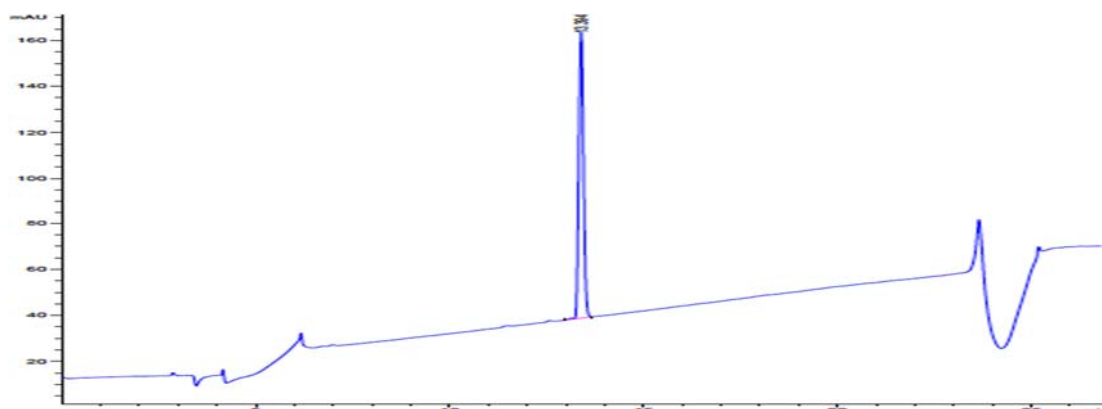
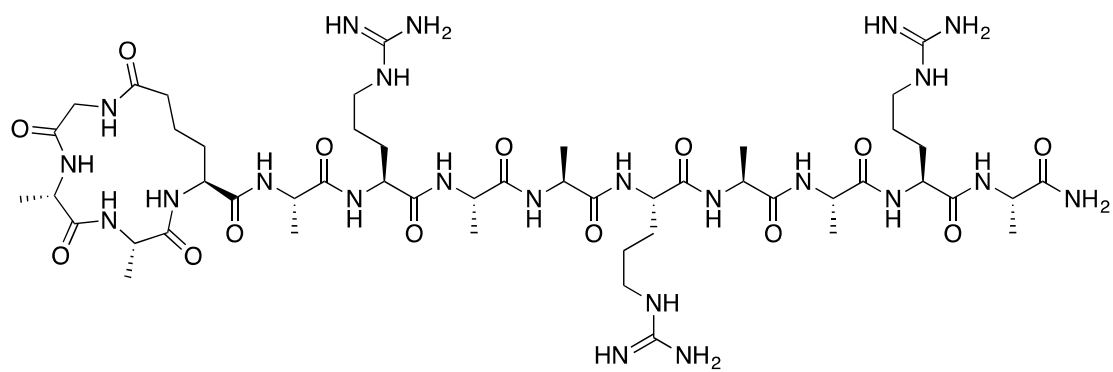
7

cyclo-(1,4)-[ARAhE]ARAARA-NH₂

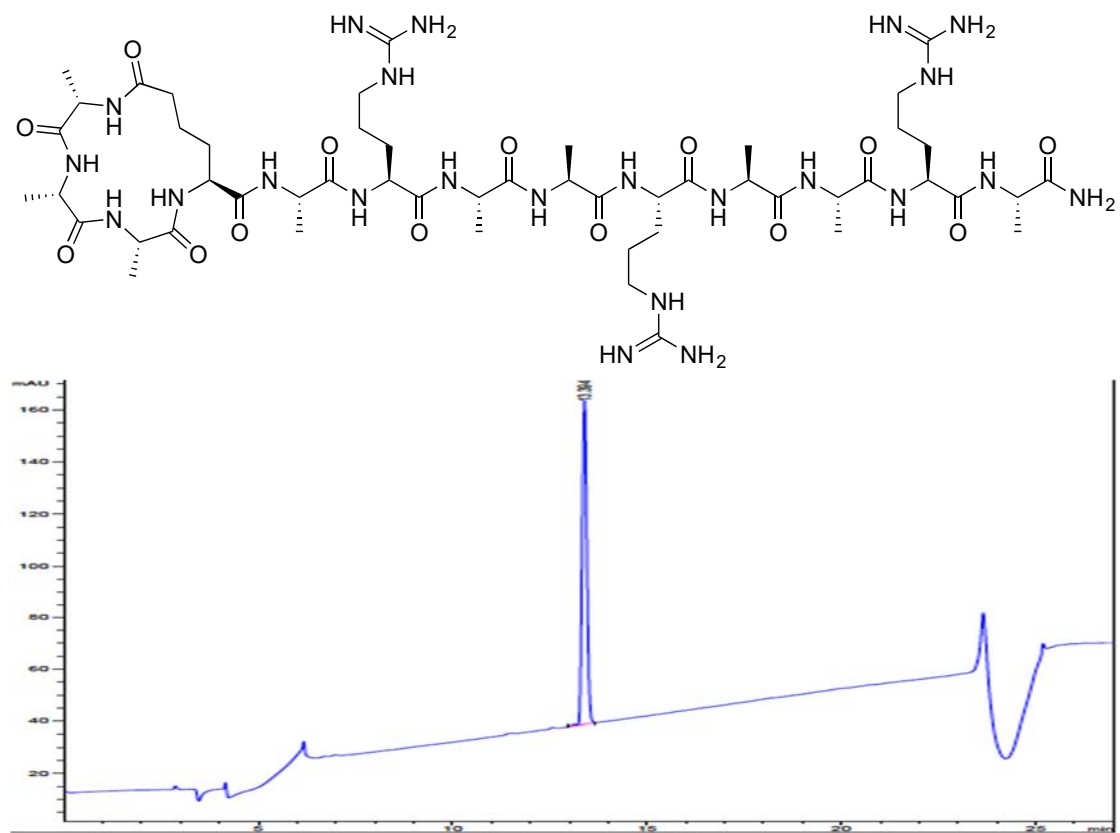
8

Ac-ARA^hEARAARA-NH₂

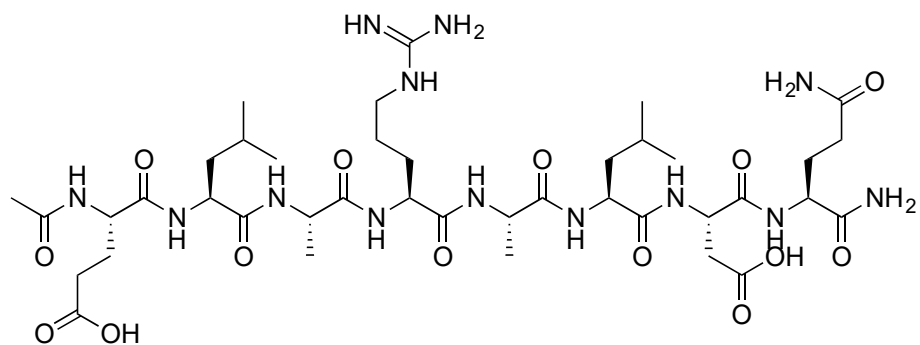
9

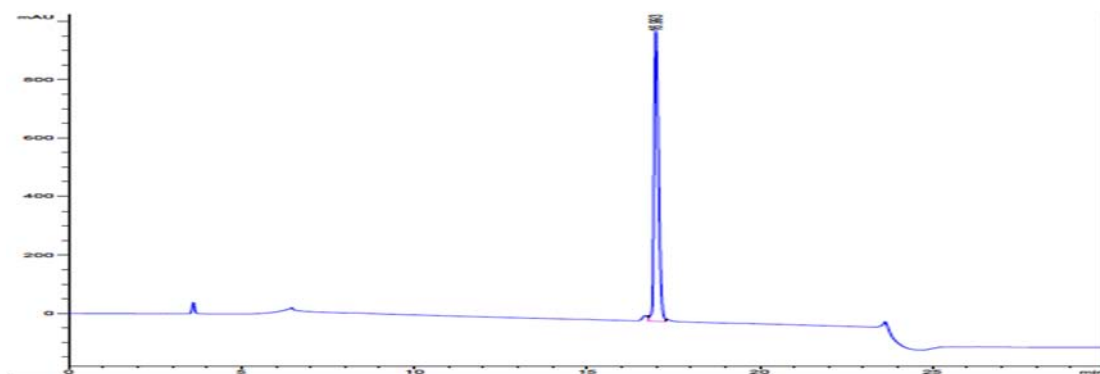
cyclo-(1,4)-[GAA^hE]ARAARA-NH₂

10

cyclo-(1,4)-[AAAhE]ARAARAARA-NH₂

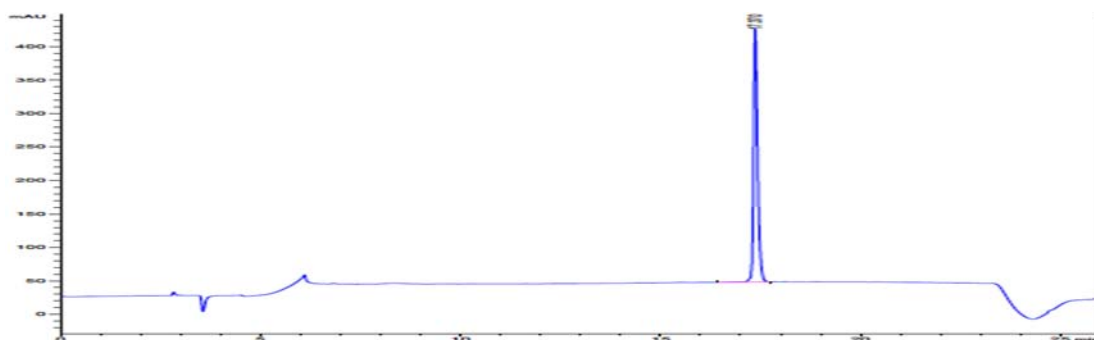
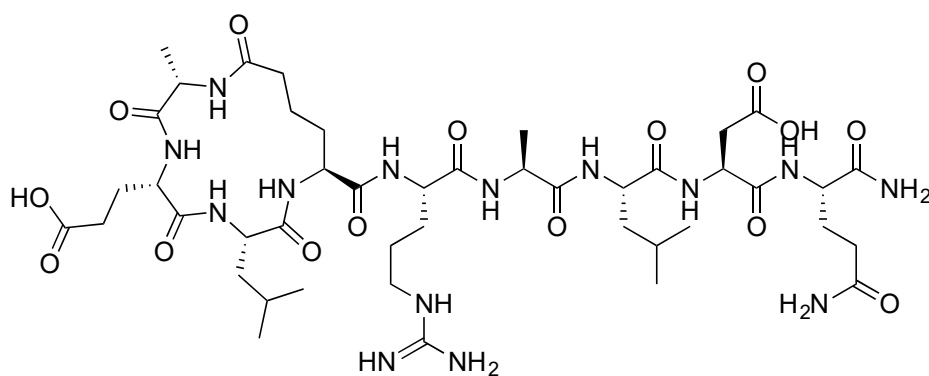
11

Ac-ELGRALDQ-NH₂



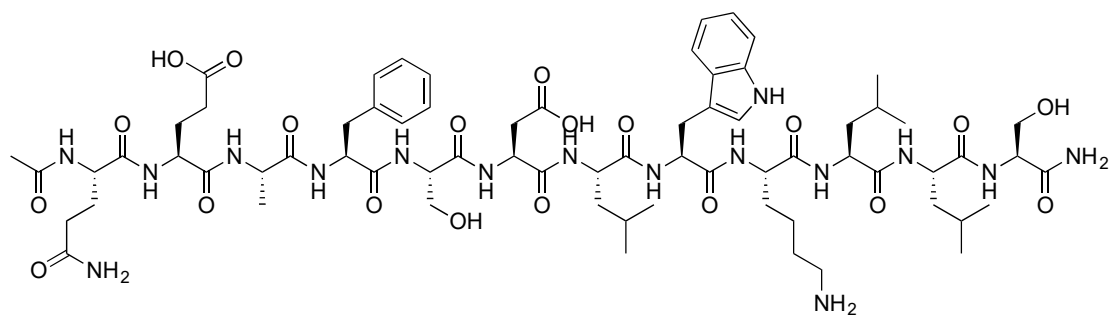
12

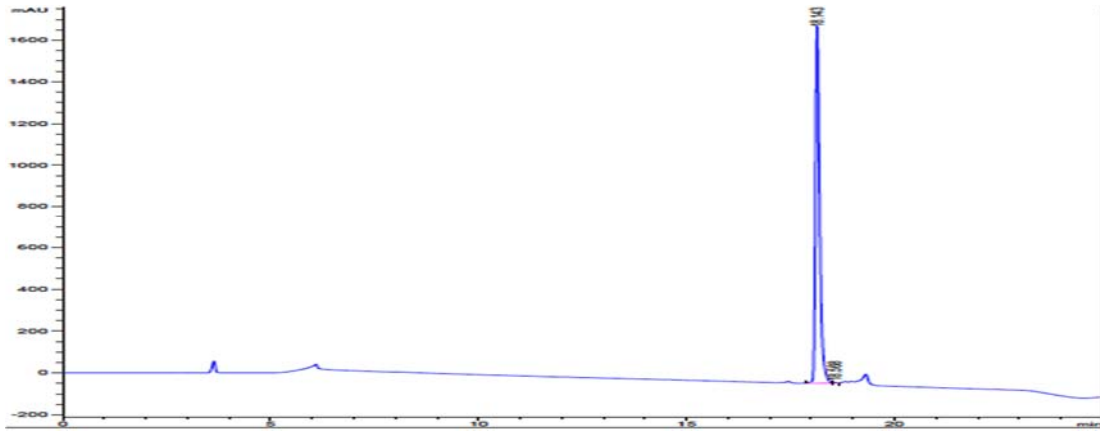
cyclo-(1,4)-[AELhE]RALDQ-NH₂



13

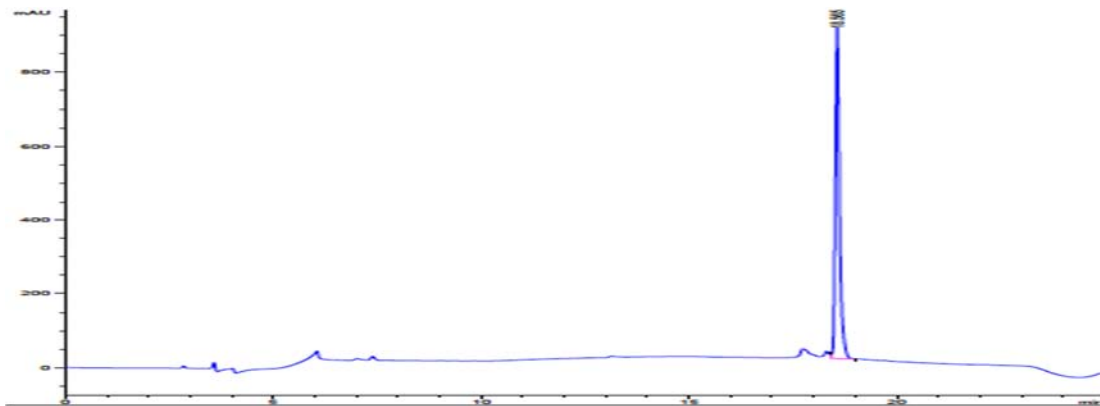
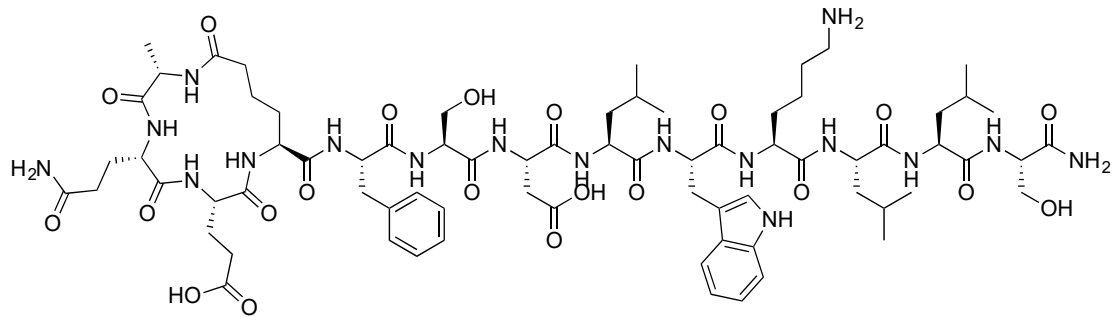
Ac-QEAFSDLWKLLS-NH₂





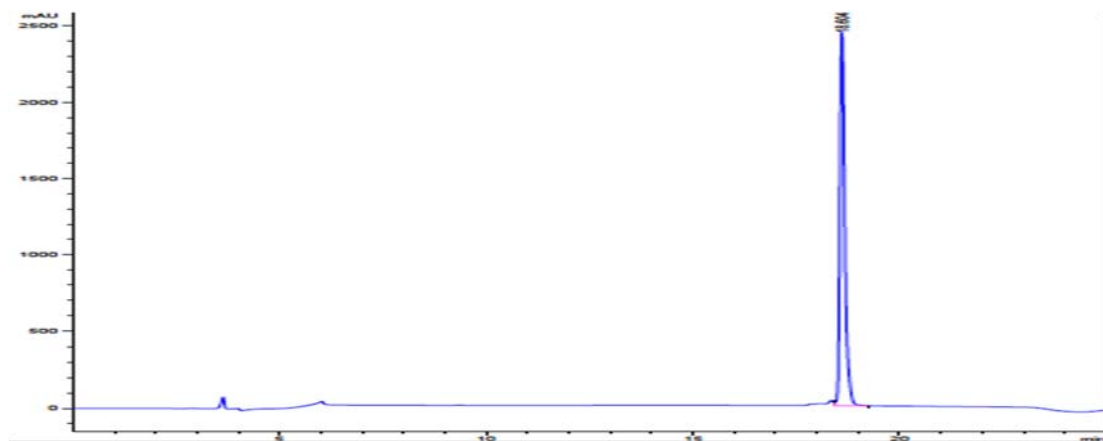
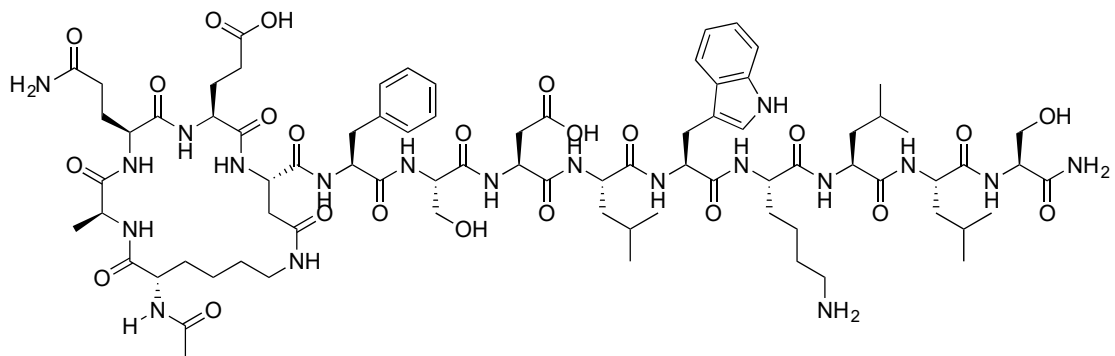
14

cyclo-(1,4)-[AQEhE]FSDLWKLLS-NH₂



15

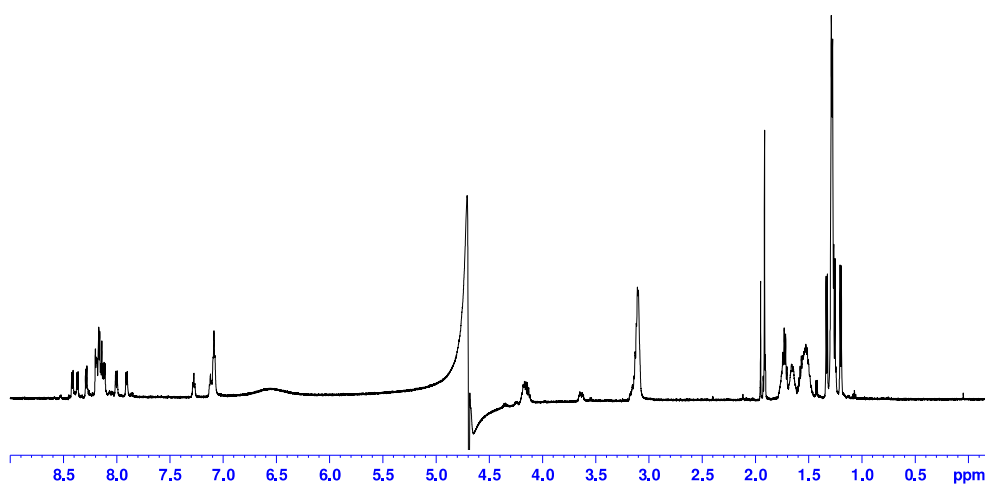
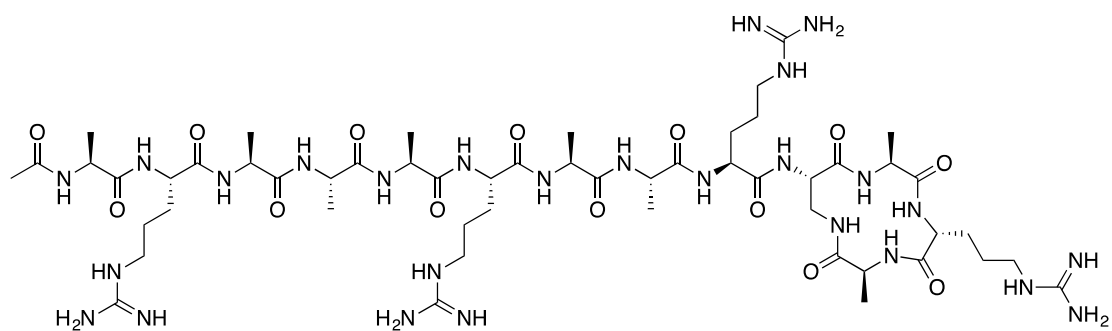
Ac-cyclo-(1,5)-[KAQED]FSDLWK-LLS-NH₂



NMR Spectrum ^1H

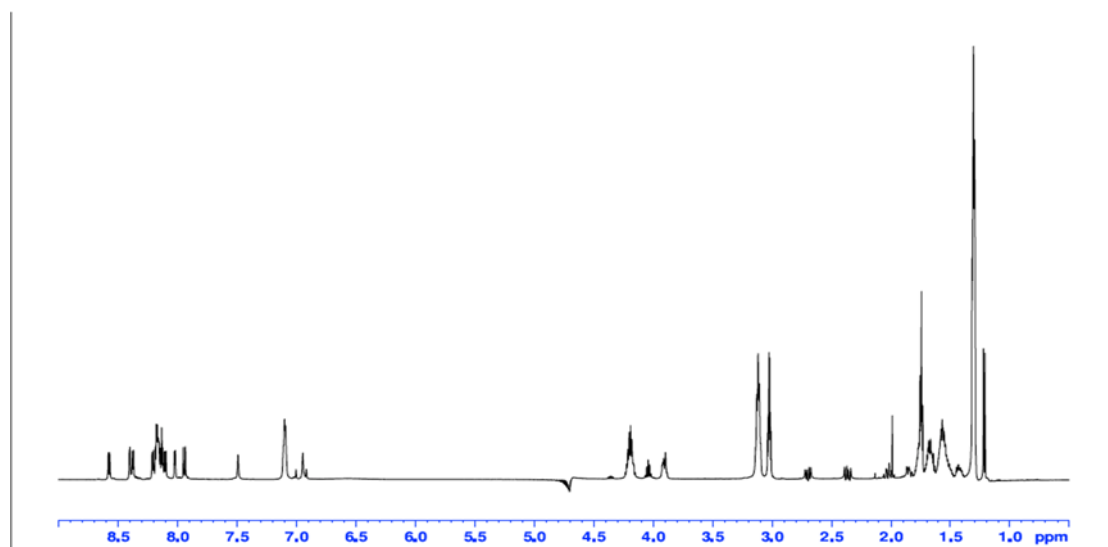
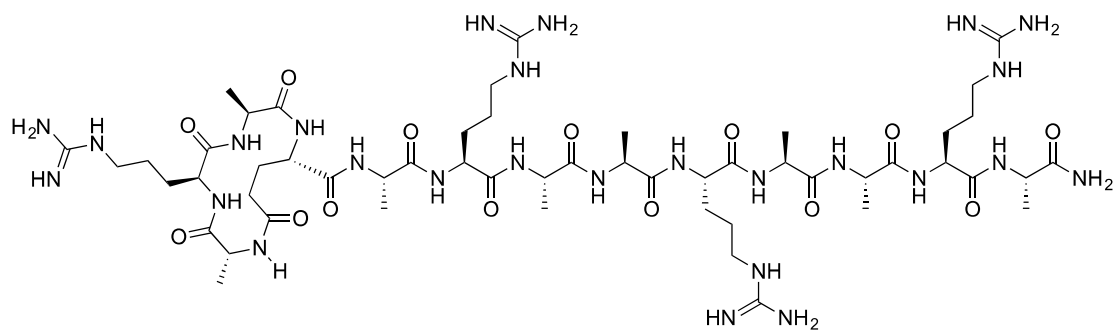
5

Ac-cyclo-(10, 13)-ARAARAARA[DapARA]



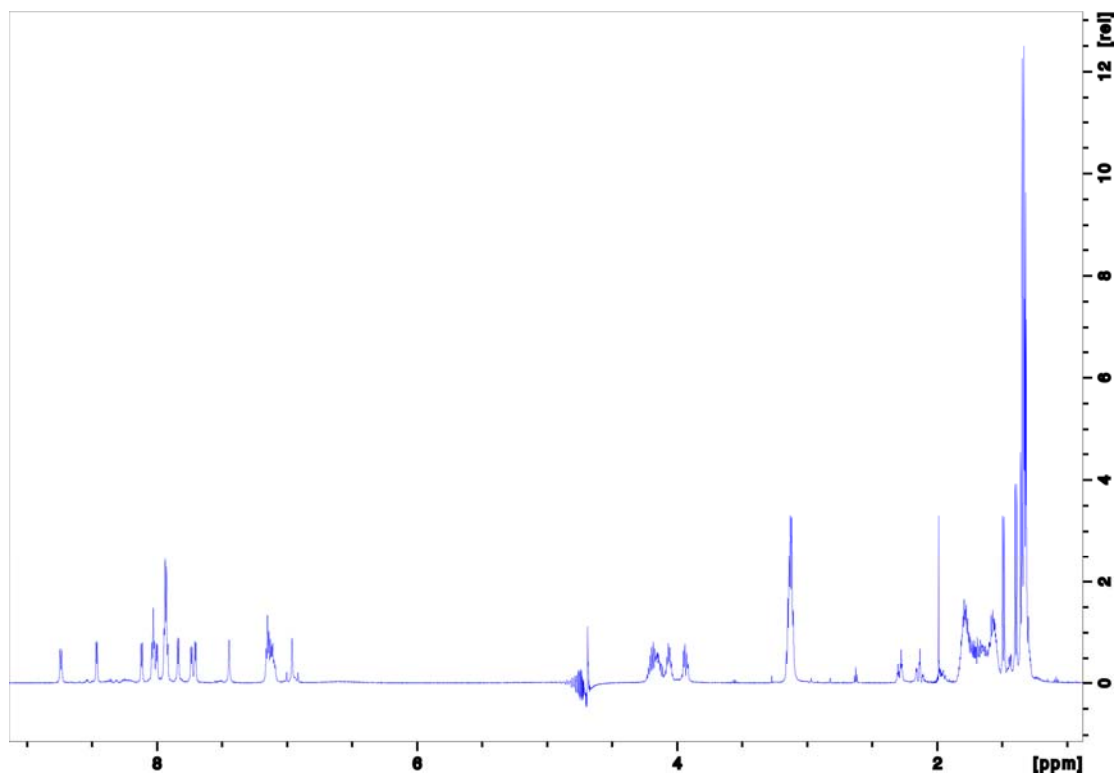
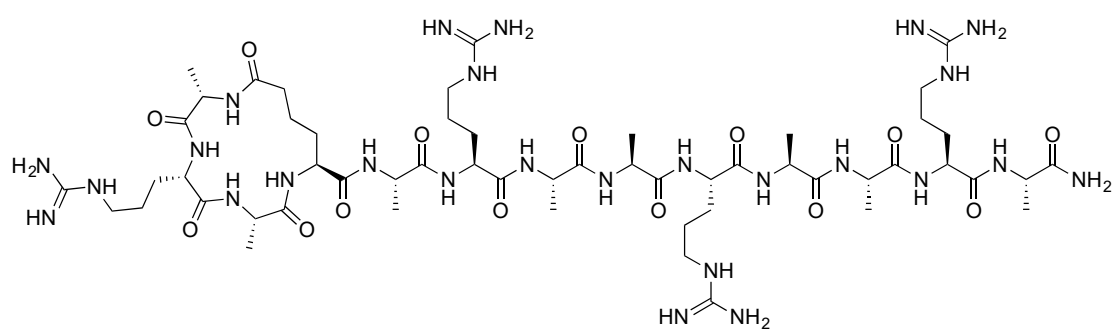
6

cyclo-(1,4)-[ARAE]ARAARA-NH₂



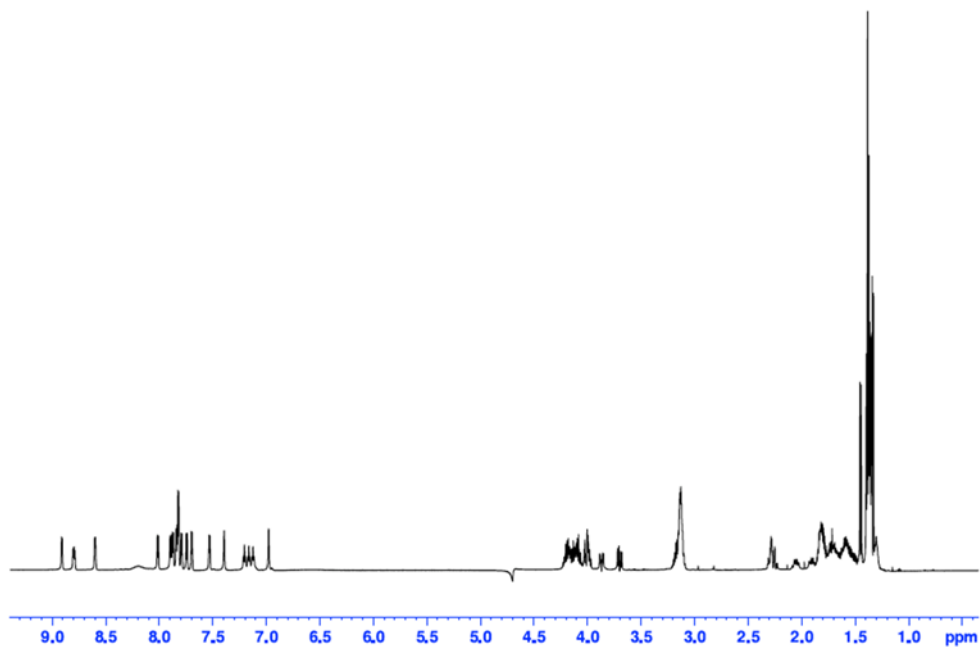
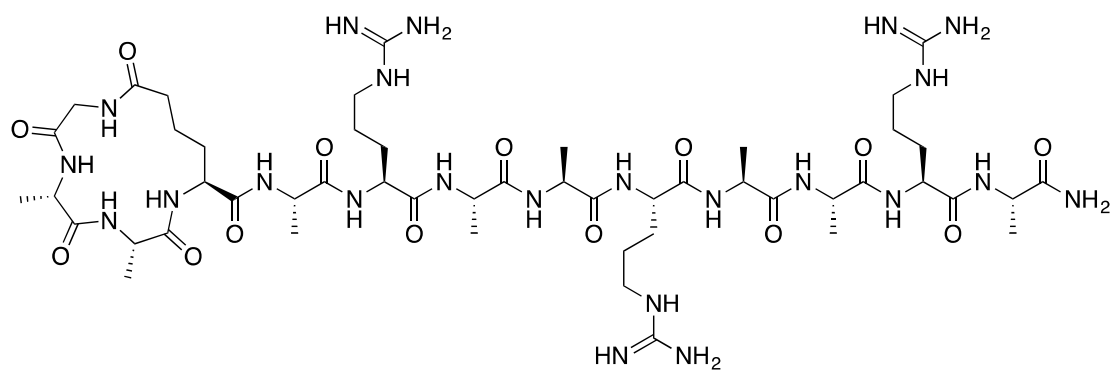
7

cyclo-(1,4)-[ARAhE]ARAARAARA-NH₂



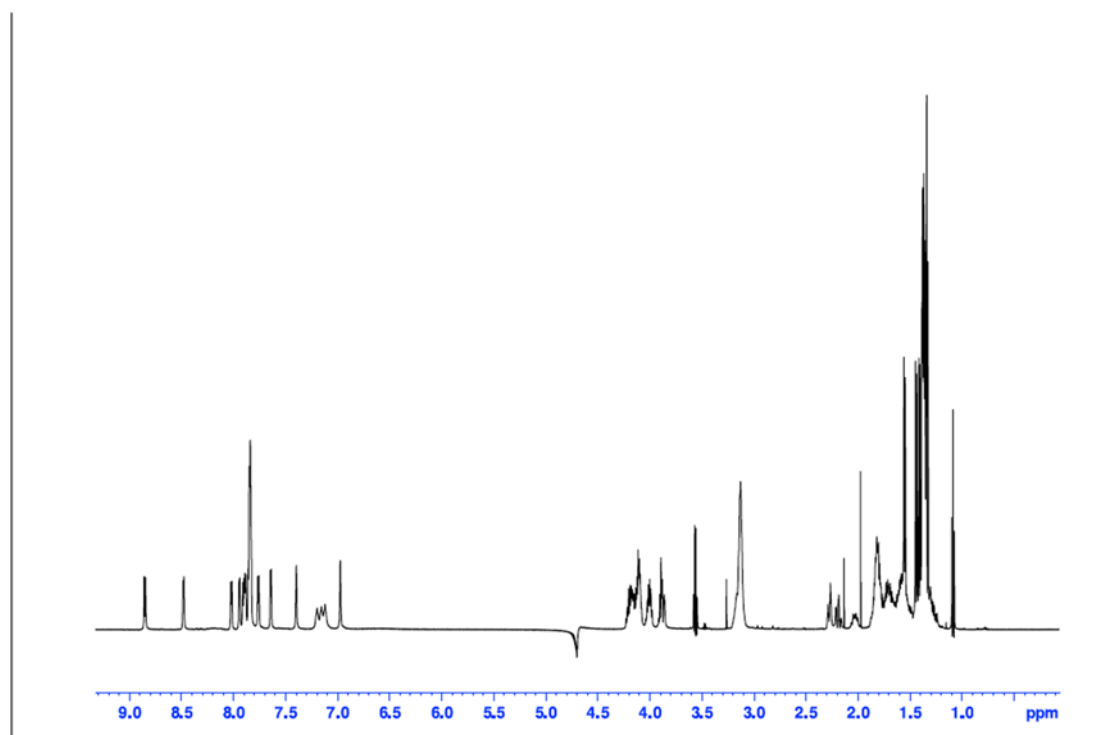
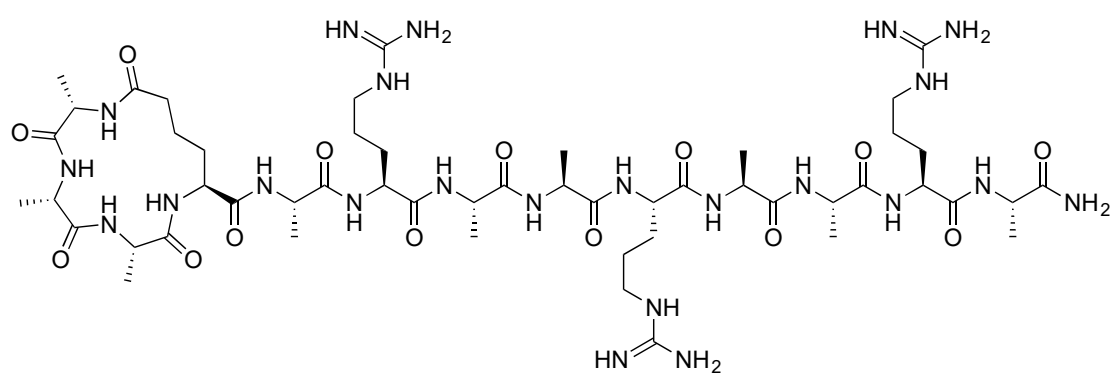
9

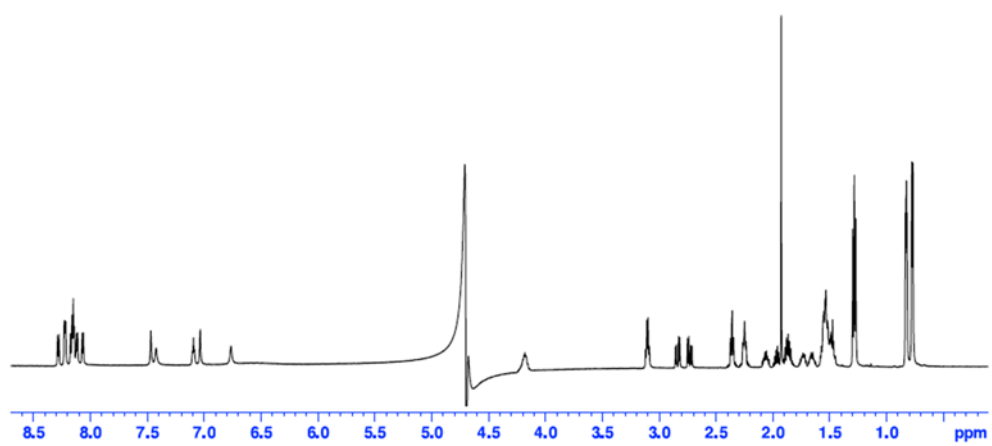
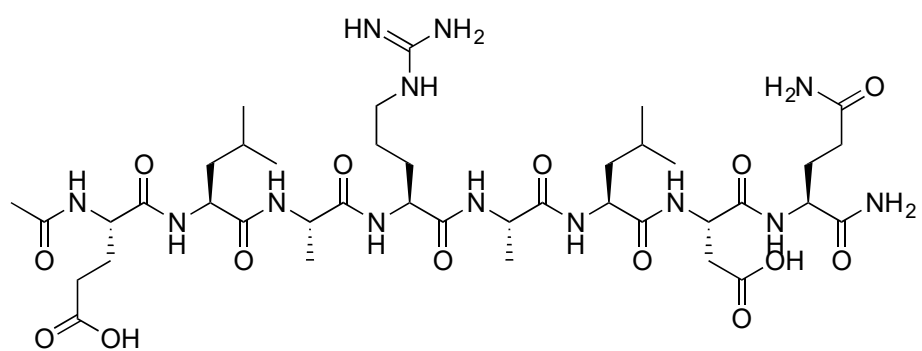
cyclo-(1,4)-[GAAhE]ARAARAARA-NH₂

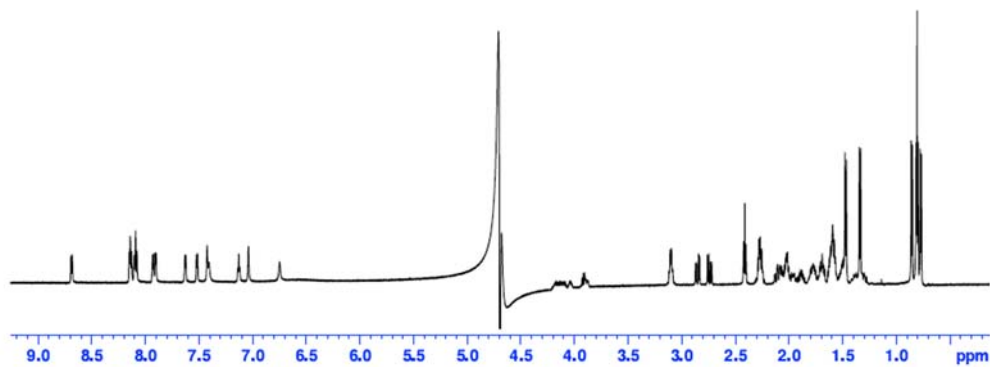
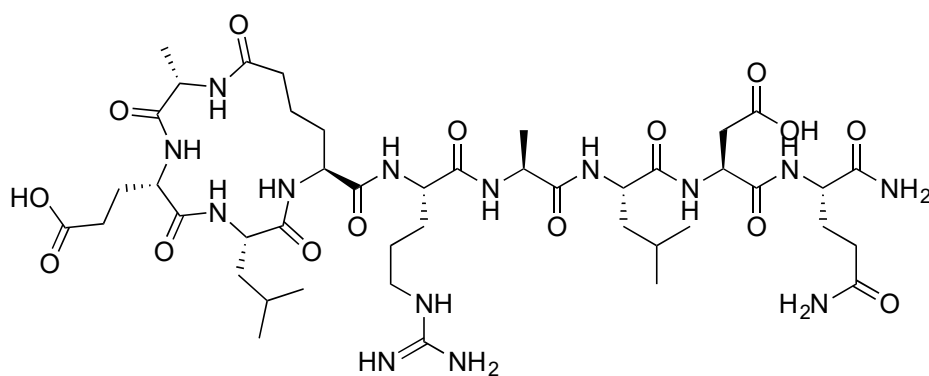


10

cyclo-(1,4)-[AAAhE]ARAARA-NH₂

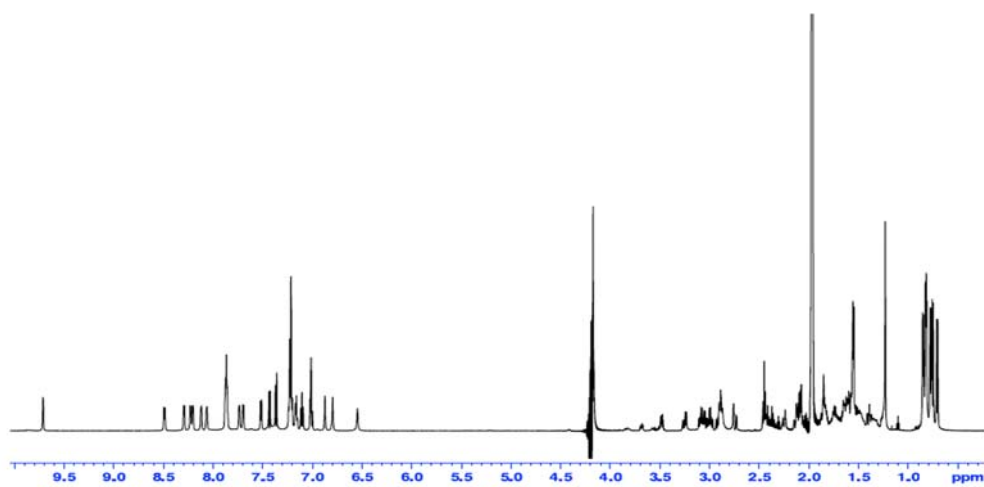
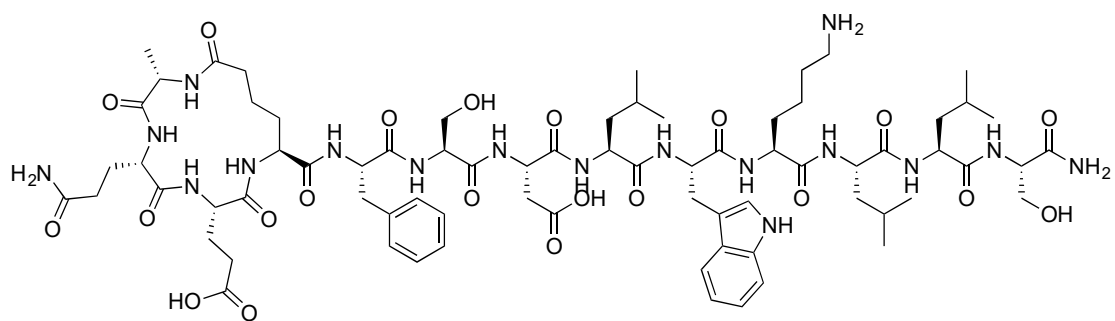


Ac-ELARALDQ-NH₂



14

cyclo-(1,4)-[AQEhE]FSDLWKLLS-NH₂



15

Ac-cyclo-(1,5)-[KAQED]FSDLWK-LLS-NH₂

